

Analysis of the Interaction between Human Operator and Automated Dispatch in Haul Truck Scheduling

Patrick Stahl, Birsen Donmez, Greg A. Jamieson

University of Toronto, Department of Mechanical and Industrial Engineering, Toronto, ON, Canada

This paper presents the findings from a field study of human-automation interaction in an open pit gold mine. Motivated by an earlier study that identified problematic interaction between haul truck operators and dispatch interfaces, focus groups and questionnaires were used to understand what causes the general attitude of suspicion towards the system. Overall trust in the dispatch interface, as well as usability and functionality of the system, were judged as slightly positive. However, the inability of the system to react efficiently to sudden changes on site results in operator frustration. We argue that consequently, the human operator should be utilized as a sensor by the dispatch system. Through operator involvement in the stage of information acquisition, the system's response to sudden changes can be improved, and discontent reduced.

INTRODUCTION

In open pit mining, an essential step in the extraction of raw materials is the transportation of the material from the place of excavation to the next processing site. Haul trucks are the primary means of transportation, and the efficiency of the haul truck fleet is of utter importance to the overall productivity of the mine. One important indicator for the productivity of a mine is the amount of material hauled per shift. Accordingly, an important goal is to ensure that trucks serve shovels continuously to minimize waiting time (miners talk about a 'hanging shovel').

On sites with multiple shovels and dumps, ensuring efficient haul traffic requires complex planning and scheduling. To this end, mines have introduced automated dispatch systems. Every truck is equipped with a touchscreen dispatch interface, and an underlying algorithm issues assignments mostly independent of the truck operator. However, the interaction between the dispatch system and the truck operator has not received much attention. This is characteristic of an industry that has put major emphasis on developing technical and hardware solutions, while paying only limited attention to the interaction between the operator and new technology (Munro & Tilyard, 2009). A recent study conducted by the authors in two open pit gold mines found that operators perceived the dispatch system to be inflexible and hard to operate, with the result being a general feeling of discontent (Stahl, Donmez, & Jamieson, 2011). We argued that the general operator discontent with respect to dispatching stemmed largely from the high level of automation, such that operators felt out of the loop and even became complacent in some cases. Moreover, the dispatch system did not react automatically to unanticipated and abrupt changes on site, creating situations in which operators had to manually request reassignments, and thereby amplifying discontent with the system.

To analyze this operator discontent and the troubled operator-automation interaction we undertook a field study in an open pit mine. Through the use of focus groups, questionnaires and observation, we established an

understanding of the general operator perspective. This operator perspective will be presented, and recommendations will be given as to how to improve operator-dispatch interaction.

METHOD

Field Study Environment

In June and July of 2011, we undertook a field study at a Nevada open pit gold mine for a total of four weeks. In this period, we spent 20 days on site observing and gathering data in the form of audio recordings, video recordings, photos, and handwritten notes.

The mine operated with 150 haul truck operators divided across four crews. Shifts were divided evenly across these four crews, and each operator was part of only one crew. Trucks operated on an extensive road network and served up to four shovels operating at the same time. Consequently, the site was characteristic of a relatively high traffic volume mine. Complex and dynamic dispatching algorithms were employed to optimize operational efficiency and productivity.

In addition to observing drivers at work, the study included focus group sessions, as well as questionnaires. Four focus groups were held with five operators each (always from a single crew), and lasted between 45 and 60 minutes. Focus groups were designed in accordance with the literature (Onwuegbuzie, Dickinson, Leech, & Zoran, 2009), with the relatively low number of five participants attributed to limitations of the work schedule on site. The sessions were structured in two parts. An initial ten-minute presentation by the researcher introduced the project, laid out expectations, and attempted to sensitize the participants to the general topic of human – automation interaction. An open discussion followed, in which participants were encouraged to describe and evaluate their interaction with the dispatch interface, as well as comment on each other's observations. The role of the researcher in this phase was passive in nature –serving only as a general guide to the discussion and making sure that all participants were included.

While we had attempted to hold all focus groups sessions before designing the questionnaire, due to time constraints and operator availability only two focus groups were conducted before the questionnaire was finalized. Questionnaires were distributed at the end of the weekly crew meeting held before the start of the shift on four consecutive Tuesdays. Altogether, 86 questionnaires were returned. Table 1 presents the distribution of questionnaire respondents across crew and gender.

Table 1: Number of Questionnaire Respondents across Crew and Gender (N.S. = Not Specified)

	Crew A	Crew B	Crew C	Crew D
Female	2	0	3	2
Male	24	9	25	19
Not specified	1	1	0	0
Total	27	10	28	21

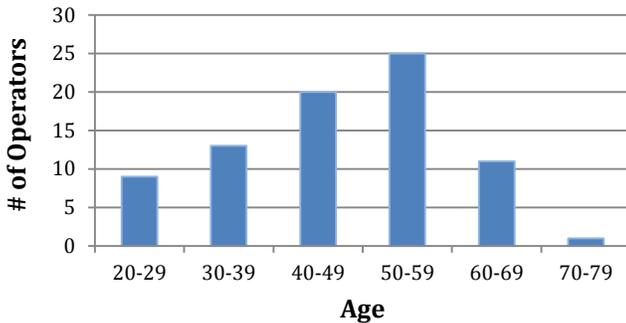


Figure 1: Age Distribution of Questionnaire Respondents

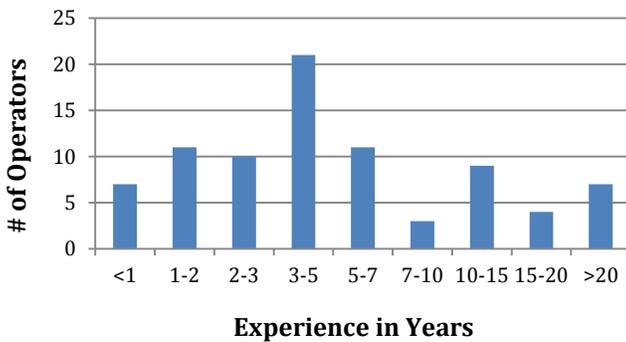


Figure 2: Experience Distribution of Questionnaire Respondents

The ages of respondents ranged from 21 to 72 years, with a mean (M) of 47 and a standard deviation (SD) of 12.09. There was a linear increase in the number of drivers across age categories up to the age of 59 (Figure 1). Haul truck experience peaked at 3-5 years for the sample, but exceeded 20 years in five cases (Figure 2). The average operator experience was approximately 7.5 years (SD = 7.18). Focus group participants were selected from the same population as

questionnaire participants, and effort was undertaken to select in a way that was representative of the group’s demographics.

Questionnaire Composition

The first six questionnaire items addressed operator trust in the automated system, and were inspired by existing trust-in-automation research (Jiun-Yin, Bisantz, & Drury, 2000). Ten more questions addressed usability and functionality of the interface. The remaining 13 questions were specific to issues that had arisen during the initial, ethnographic study conducted by the authors (Stahl et al., 2011), during observations of mine personnel performed in the current field study, and most importantly, during the focus group discussions. Topics included the modality of communication, reliability and extent of information communicated through the interface, and the extent of operator involvement.

Finally, three items posed different alternatives for the kind of information operators felt should be included in the dispatch interface (such as “best route to next assignment” or “closed roads”), for the information they would like to be able to input into the system (such as “changes in road condition” or shovel status”), and finally for the modality of system alerts (“auditory”, “visual”, or “both”). A final item allowed for additional comments.

Participants indicated their level of agreement or disagreement with 29 statements using a seven-point Likert scale (Likert, 1932). Twenty-four of the statements used the same attributes for the seven values – strongly disagree (1), disagree (2), slightly disagree (3), neutral (4), slightly agree (5), agree (6), and strongly agree (7). The wording of the attributes for five statements was modified to match the statement.

RESULTS

Focus Groups

Participants in the focus groups were generally receptive to the issue of human-automation interaction, such that highly active discussions took place with minimal guidance. Operator opinion of the dispatch system often appeared to be relatively positive initially, but a more critical attitude tended to emerge in the course of the discussions. It appeared that the operators were accepting of a system they had little influence over, but provided a wealth of constructive criticism about the system once they were encouraged to think critically. Recurring topics across crews were:

- The lack of consideration from the dispatch system for information that operators perceive as important, such as road and weather conditions, shovel status, and load count.
- The inability of the system to react to sudden changes on site, such as changes in shovel status or blocked roads.
- The non-optimal implementation of the manual reassignment option: confirmations often take too long, the system sometimes has to be manually updated in order to display new assignments.

- A general feeling of being detached from the dispatch interface and frequently disagreeing with assignments.

Questionnaires

Participant responses to trust-related questions were averaged to create an overall trust rating. Similarly, responses to usability and functionality questions were also averaged to create overall usability and functionality ratings, respectively. Given that the dependent variables were generated through the averaging of Likert-scale data, the data were no longer discrete and thus were analyzed with statistical methods appropriate for continuous data. Analyses were conducted on these three dependent variables (i.e., trust, usability, and functionality ratings) to investigate effects of gender (male, female), age (<30, 30 - 59, >59), experience (<=2, >2 - <10, >= 10), and crew (A,B,C,D). The crew effect was tested because of the pairing of crews with specific dispatch controllers – if one crew were to show higher trust and functionality ratings, for example, this may be due to a particularly skilled dispatcher.

Separate statistical models were built for each predictor variable: independent t-tests for gender and separate ANOVAs for age, experience, and crew. It was not possible to investigate interaction effects due to a very small number of observations within particular design cells (e.g., there were no young female drivers; nor were there any young drivers in crew B). Model assumption checks were conducted through residual analysis.

The evaluation of the questionnaire showed moderately positive feedback altogether. The spread across the averaged feedback ratings for the individual statements was relatively small; only two statements averaged below neutral, and only one statement exceeded a rating of six.

Trust Ratings. Overall, the evaluation of trust-related statements showed moderately positive feedback, with an average of 5.05 (which corresponds to the attribute “slightly agree”) and a standard error (SE) of 0.19. Grouping by crew did not appear to have a significant effect on trust, $F(3, 81) = 0.47, p = 0.71$, nor did grouping by age or experience, $F(2, 76) = 0.20, p = 0.82$; $F(2, 80) = 0.50, p = 0.61$. The only significant finding was for gender; females on average trusted the interface less ($M = 4.19, SE = 0.18$) than males ($M = 5.06, SE = 0.20$), $t(9.29) = -3.22, p = 0.01$.

Usability Ratings. The cluster of usability statements showed similar trends. The overall average rating was 5.16 ($SE = 0.43$), again marking a moderately positive Likert value. Feedback with respect to the ease of use and the menu structure were quite positive ($M = 5.64$ and $M = 5.94$ respectively). However the position of the interface in the cockpit ($M = 5.05$) was rated less positively, and the statement about the distractive effect of the interface ($M = 3.99$) received only neutral feedback.

With respect to the predictor variables, again no effect of crew was found, $F(3, 82) = 0.41, p = 0.75$. There was a significant effect of experience, $F(2, 80) = 3.08, p = 0.05$. A subsequent Tukey post-hoc test showed a significant difference between the groups with limited and extensive experience, with more experienced operators rating the

interface as less usable (mean difference = 0.71, $lwr = 0.01$, $upr = 1.40, p = 0.047$).

The influence of age was also significant, $F(2, 76) = 3.78, p = 0.03$. Post-hoc tests identified young operators as giving better usability ratings than both, middle aged (mean difference = 0.82, $lwr = 0.06, upr = 1.57, p = 0.03$) as well as older operators (mean difference = 0.97, $lwr = 0.04, upr = 1.90, p = 0.04$). No significant difference was found between middle aged and older drivers (mean difference = 0.16, $lwr = -0.83, upr = 0.51, p = 0.84$). Finally, the initial gender dependency suggested by the lower female ($M = 4.43, SE = 0.39$) than male mean ($M = 5.19, SE = 0.43$) was only marginally significant, $t(7.32) = -2.26, p = 0.06$.

Functionality Ratings. Questionnaire items on functionality had operators indicate whether the system provided the information they needed ($M = 5.52$), whether it saved them time ($M = 5.27$), whether the system worked well ($M = 4.68$), was consistent ($M = 4.75$), and whether it was considered useful ($M = 5.55$). The overall average for this group of items was 5.15 ($SE = 0.19$).

Females gave significantly lower functionality ratings ($M = 4.17, SE = 0.35$) than males ($M = 5.23, SE = 0.18$), $t(10.79) = -4.42, p < 0.01$. No differences were observed for age, experience or crew, $F(2, 76) = 0.70, p = 0.50$; $F(2, 80) = 0.77, p = 0.47$; $F(3, 81) = 0.32, p = 0.80$.

Table 2 summarizes the statistical significance of the four predictor variables with respect to trust, usability and functionality ratings.

Table 2: Significance of Predictor Variables on Trust, Usability, and Functionality Ratings

	Gender	Age	Experience	Crew
Trust	p < 0.05	p > 0.05	p > 0.05	p > 0.05
Usability	p > 0.05	p < 0.05	p = 0.05	p > 0.05
Functionality	p < 0.01	p > 0.05	p > 0.05	p > 0.05

Problem-Specific Ratings: A look at the specific questions (Figure 3) showed that operators largely agreed with the assignments they received through the system ($M = 5.27$), but that specifically in situations where adjustments needed to be made as a consequence of sudden, unexpected changes, ratings became relatively critical ($M = 3.86$). Asked about whether operators would welcome the opportunity to input such changes, participants reacted positively ($M = 5.20$). Further, participants already used existent system features to indicate their status (such as lunch breaks, or fuel breaks) ($M = 5.45$), and would appreciate the system to take into account additional information, such as weather and road condition ($M = 5.10$). Operators were neutral to the idea of including GPS navigation in the interface ($M = 4.28$).

Of particular interest was the highly positive feedback with respect to the opportunity for operators to manually reassign themselves ($M = 6.31$). A neutral rating was given when asked whether the need to update the system after requests for manual reassignments (several operators had commented on the need to repeatedly force-update their interface in order for the changes to take effect) discouraged operators from using this feature ($M = 4.01$).

DISCUSSION

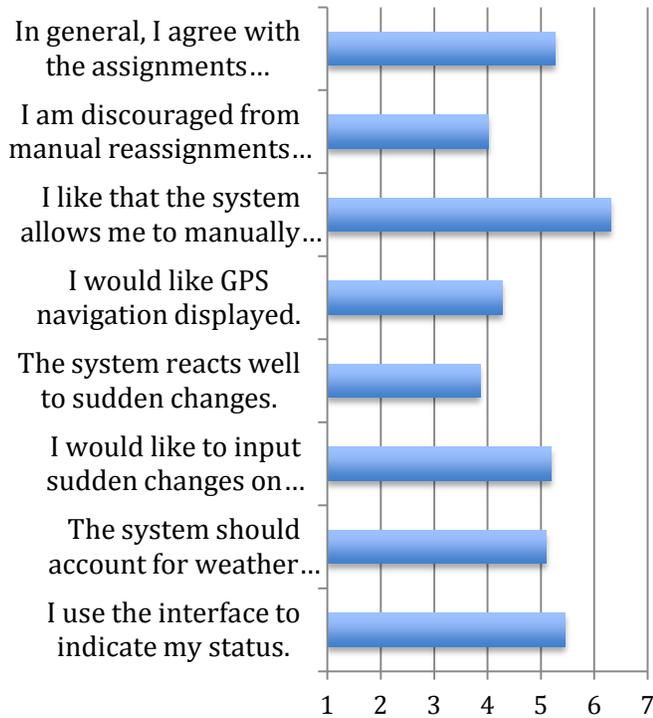


Figure 3: Mean Responses to Problem-Specific Questions

In five more statements, operators reported that they found the process to manually reassign themselves to be rather easy ($M = 5.38$), that they thought there were sometimes better assignments possible than those issued by the system ($M = 4.28$), that they had to reassign themselves several times per shift ($M = 4.35$), that confirmations for manual reassignments took too long ($M = 5.05$), and that they judged their influence over the system to be moderate ($M = 4.10$). Confronted with the choice of modality for communication, 61 out of 86 operators stated that they would have preferred both, visual and audio alerts. 16 would have preferred an exclusively visual interface, and 9 would have chosen only auditory feedback.

Finally, two multiple choice questions asked for the type of information that operators would like to be able to input into the system, as well as the type of information the system should display. With respect to inputting information, 68 operators indicated “shovel status”, and 66 “road blocks”. 51 operators would like to input “changes in road condition”, and 40 “changes in weather”. Other operator ideas that had not been taken into account when creating the questionnaire were the reporting of emergency vehicles on site, of dusty roads, of the load count, and finally the opportunity to input preset messages to other truck drivers and/or dispatch. With respect to the type of information the interface should display, 62 operators checked the “next assignment”, 60 indicated “closed road”, and 51 wanted the “load count” displayed. “Bad weather” was indicated in 38 responses, the routing via “best route to next assignment” accounted for 36 responses, and finally “slippery ramps” for 32. Additional proposals were “shot time” (the time of the daily blast), “load weight”, the number of trucks inbound to a specific shovel, “crew tonnage”, and finally “alerts for emergency vehicles”.

In general, operator ratings indicate that the dispatch system functions in a satisfactory and reliable manner. There is moderate trust in the system, and judgments with respect to usability and functionality are also reasonable. However, it is also apparent that the system is not viewed particularly positively – overall operators “slightly agree” with all three measures.

An interesting finding is the gender dependency in the ratings – females consistently judged the system more negatively than did males, a finding that has also been reported in the field of economics (Buchan, Croson, & Solnick, 2008). To explain the influence of gender on functionality in particular, but also on usability, future studies should be designed to allow for the investigation of interaction effects between trust, usability, and functionality. Arguing on the basis of the conceptual model of the processes governing trust (Lee & See, 2004), both functionality and usability are likely to influence belief formation and information assimilation with respect to the system, thereby directly influencing trust.

With respect to the usability of the system, younger drivers appeared to rate the system more positively. The reason for this tendency is very likely a generally higher proficiency with computer-based, menu driven interfaces (Veen & Vrakking, 2006). While the younger driver generations of up to 30 years of age are highly likely to have grown up with computer applications, older drivers are less likely to be proficient in the use of menu driven touchscreen interfaces.

Of particular interest, however, are the ratings for the problem-specific questions. The most positive questionnaire responses were given for the statement “I like that the system allows me to manually reassign” myself, and the lowest ratings were given in response to “The system reacts well to sudden changes”. These results support the general impression - gained from observations and focus groups – that the dispatch system is highly automated, acts on a long-term, global level, and is not sensitive to sudden changes on site. While operators naturally cannot judge from the global dispatch perspective, they are directly involved in and aware of short-term changes on site that can impact their assignments, up to the point of making their assignments unfeasible or impossible. Such situations were witnessed frequently during observations on site, and operators pointed them out in focus groups and questionnaires. An example of a sudden change on site that the dispatch system does not account for is the shovel status. A shovel has to move frequently to reposition itself relative to a gradually progressing bank. Usually a bulldozer will move in front of the shovel (to clean up rocks in front of the shovel). Altogether the moving process can take several minutes. Since the dispatch system does not take this downtime of the shovel into account, the result is often a long line up of idling trucks, and truck operators ask for manual reassignments. A second example is that road conditions frequently change to the point that operators temporarily do not use a particular part of road, often due to extensive rockspill (excavated material that has

rolled of the bed of a truck and is a potential threat for traffic safety) or very wet, slippery ground.

From the operator perspective, situations like these are viewed as failures of the automated system, thereby lowering the perceived reliability. The reason for trust ratings to be relatively unaffected by this, contrary to the suggestions of existing research (Lee & See, 2004), may be the general, passive acceptance operators treat the system with – they cannot change it and have grown used to it being unable to deal with sudden changes on site. These “failures” are expected, and operators are accustomed to having to compensate for them.

However, the operator is also an under-used asset with respect to the gathering of information relevant for dispatching decisions. The willingness of operators to actively input information into the system was apparent in the questionnaire results, and further supports the objective of including the operator in the dispatch system (see Stahl et al., 2011).

Arguing in terms of the model of stages and levels of automation (Parasuraman, Sheridan, & Wickens, 2000), it is specifically the stage of information acquisition in which the level of automation should be adjusted to make use of the free information available through equipment operators. It should be noted that this adjustment would effectively lower the level of automation, even though the functionality of the system would likely be improved. This strategy would be contrary to beliefs about higher levels of automation to be favorable, or about higher levels of automation to be desirable specifically in information acquisition (Inagaki, 2003). The argument made is therefore to go beyond human machine interaction and levels of automation, and into human machine cooperation (Hoc, 2000). Also, by permitting the operator to input dispatch-relevant information into the system, operator involvement could be increased, and therefore counteract the perception of being detached from the system.

Finally, while the impressions from observations, focus groups, and questionnaires were generally in consensus, there was a distinct tendency for questionnaire results to be less critical. Where focus group discussions showed operators to frequently disagree with the assignments given, questionnaires reported general agreement with the assignments. Similarly, discussions revealed operators desiring more task-specific information, while questionnaire feedback suggested that the system communicated all required information. This tendency may be due in particular to the varying contextual sensitivity of operators. While the focus groups encouraged them to think critically about the issues presented, the questionnaires were answered at the end of the weekly crew meeting and without a prior introduction of the issue at hand. In order to elicit critical opinion from questionnaire participants being asked to judge any practice that has become highly habitual to them, it may be necessary to sensitize them to the issue at hand.

CONCLUSION

In order to assess haul-truck operator opinion with respect to their interaction with an automatic dispatch system, we conducted a study that used focus groups and questionnaires. Ratings with respect to trust, usability and functionality were

moderately positive. Trust and functionality ratings proved to be dependent on gender (with females giving more critical ratings than males), while usability ratings were dependent on both operator age and experience (with young, less experienced operators giving better ratings).

One specific problem identified with the dispatch system was the lack of sensitivity to the operator perspective, specifically in terms of not being able to react to sudden changes on site that required operators to deviate from their assignments. It was suggested to allow operators more influence over the system, and in particular to allow them to actively input information that is relevant to dispatching.

For future work, the impact of different levels of operator inclusion into information acquisition on trust and satisfaction with the system, but also on haul efficiency, will be investigated in the context of a simulator study. We want to focus on optimizing productivity and safety by means of altering human-automation interaction.

References

- Buchan, N. R., Croson, R. T. A., & Solnick, S. (2008). Trust and Gender: An Examination of Behaviour and Beliefs in the Investment Game. *Journal of Economic Behaviour & Organization*, 68(3-4), 466–476.
- Hoc, J.-M. (2000). From Human - machine interaction to human - machine cooperation. *Ergonomics*, 43(7), 833–843.
- Inagaki, T. (2003). Adaptive Automation: Sharing and Trading of Control. *Handbook of Cognitive Task Design* (S. 147–169). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Jiun-Yin, J., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an Empirically Determined Scale for Trust in Automated Systems. *International Journal of Cognitive Ergonomics*, 4(1), 53–71.
- Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, 46(1), 50–80.
- Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, 140, 44–53.
- Munro, P. D., & Tilyard, P. A. (2009). Back to the Future - Why Change Doesn't Necessarily Mean Progress. *Conference Proceedings from the Tenth Mill Operators' Conference 2009*. Tenth Mill Operators' Conference 2009, Adelaide, South Australia.
- Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., & Zoran, A. G. (2009). A Qualitative Framework for Collecting and Analyzing Data in Focus Group Research. *International Journal of Qualitative Methods*, 8(3), 1–21.
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A Model for Types and Levels of Human Interaction with Automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3), 286–297.
- Stahl, P., Donmez, B., & Jamieson, G. (2011). A Field Study of Haul Truck Operations in Open Pit Mines. *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting* (S. 1845–1849). 55th Annual Meeting of the Human Factors and Ergonomics Society, Las Vegas, NV.
- Veen, W., & Vrakking, B. (2006). *Homo Zappiens: Growing Up in a Digital Age*. London: Network Continuum Education.